

Harsch, Claudia; Schröder, Konrad

## **Textrekonstruktion: C-Test**

*Beck, Bärbel [Hrsg.]; Klieme, Eckhard [Hrsg.]: Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistung International). Weinheim u.a. : Beltz 2007, S. 212-225*



Quellenangabe/ Reference:

Harsch, Claudia; Schröder, Konrad: Textrekonstruktion: C-Test - In: Beck, Bärbel [Hrsg.]; Klieme, Eckhard [Hrsg.]: Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistung International). Weinheim u.a. : Beltz 2007, S. 212-225 - URN: urn:nbn:de:0111-opus-32578 - DOI: 10.25656/01:3257

<https://nbn-resolving.org/urn:nbn:de:0111-opus-32578>

<https://doi.org/10.25656/01:3257>

in Kooperation mit / in cooperation with:

# **BELTZ**

<http://www.beltz.de>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

**peDOCS**  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipt.de](mailto:pedocs@dipt.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Bärbel Beck / Eckhard Klieme (Hrsg.)

# **Sprachliche Kompetenzen**

Konzepte und Messung

DESI-Studie

(Deutsch Englisch Schülerleistungen International)

Beltz Verlag · Weinheim und Basel

Dr. *Bärbel Beck* ist Diplompsychologin und Projektkoordinatorin am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt a.M.

Prof. Dr. *Eckhard Klieme* ist Direktor des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) in Frankfurt a.M.

Diese Studie wurde im Auftrag der Kultusministerkonferenz erstellt. Für die Richtigkeit des Ergebnisses der Studie trägt das »Deutsche Institut für Internationale Pädagogische Forschung« allein die Verantwortung.

Das Werk und seine Teile sind urheberrechtlich geschützt. Jede Nutzung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages. Hinweis zu § 52a UrhG: Weder das Werk noch seine Teile dürfen ohne eine solche Einwilligung eingescannt und in ein Netzwerk eingestellt werden. Dies gilt auch für Intranets von Schulen und sonstigen Bildungseinrichtungen.

Lektorat: Peter E. Kalb

© 2007 Beltz Verlag · Weinheim und Basel

[www.beltz.de](http://www.beltz.de)

Herstellung: Klaus Kaltenberg

Satz: Deutsches Institut für Internationale Pädagogische Forschung

Druck: Druckhaus »Thomas Müntzer«, Bad Langensalza

Printed in Germany

ISBN 978-3-407-25398-9

# Inhaltsverzeichnis

<i>Bärbel Beck / Eckhard Klieme</i> Einleitung.....	1
--	---

## Übergreifende Konzeptualisierung sprachlicher Kompetenzen

<i>Nina Jude / Eckhard Klieme</i> Sprachliche Kompetenz aus Sicht der pädagogisch-psychologischen Diagnostik.....	9
---	---

<i>Günter Nold / Heiner Willenberg</i> Lesefähigkeit .....	23
---	----

<i>Claudia Harsch / Astrid Neumann / Rainer Lehmann / Konrad Schröder</i> Schreibfähigkeit.....	42
--	----

<i>Wolfgang Eichler / Günter Nold</i> Sprachbewusstheit .....	63
--	----

## Messung sprachlicher Kompetenzen

<i>Johannes Hartig</i> Skalierung und Definition von Kompetenzniveaus .....	83
--	----

<i>Jürgen Rost</i> Definition von Kompetenzniveaus mit Hilfe von Mischverteilungsmodellen .....	100
---	-----

## Kompetenzmodelle und Kompetenzniveaus im Bereich des Deutschen

<i>Heiner Willenberg</i> Lesen.....	107
--	-----

<i>Heiner Willenberg / Steffen Gailberger / Michael Krelle</i> Argumentation .....	118
---	-----

<i>Heiner Willenberg</i> Wortschatz.....	130
---	-----

<i>Günther Thomé / Jens Gomolka</i> Rechtschreiben.....	140
--	-----

<i>Wolfgang Eichler</i> Sprachbewusstheit .....	147
--	-----

<i>Albert Bremerich-Vos / Rüdiger Grotjahn</i>	
Lesekompetenz und Sprachbewusstheit:	
Anmerkungen zu zwei aktuellen Debatten .....	158

## **Kompetenzmodelle und Kompetenzniveaus im Bereich des Englischen**

<i>Günter Nold / Henning Rossa</i>	
Hörverstehen .....	178

<i>Günter Nold / Henning Rossa</i>	
Leseverstehen .....	197

<i>Claudia Harsch / Konrad Schröder</i>	
Textrekonstruktion: C-Test .....	212

<i>Günter Nold / Henning Rossa</i>	
Sprachbewusstheit .....	226

<i>Günter Nold / John H. A. L. De Jong</i>	
Sprechen .....	245

<i>Hermann-Günter Hesse / Kerstin Göbel</i>	
Interkulturelle Kompetenz.....	256

<i>Günther Schneider</i>	
Auf dem Weg zu Skalen für die rezeptiven	
Kompetenzen im Bereich des Englischen .....	273

## **Ausblick**

<i>Konrad Schröder</i>	
Kompetenz, Bildungsstandards und Lehrerbildung	
aus fachdidaktischer Sicht.....	290

<i>Günter Nold</i>	
DESI im Kontext des Gemeinsamen	
Europäischen Referenzrahmens für Sprachen .....	299

<i>Sauli Takala</i>	
Relating Examinations to the Common European Framework.....	306

<i>Hermann Lange</i>	
Abschließendes Statement.....	314

Die Autorinnen und Autoren.....	318
---------------------------------	-----

*Claudia Harsch / Konrad Schröder*

## **Textrekonstruktion: C-Test**

Im DESI-Projekt werden in je eigenen Modulen verschiedene Teilkompetenzen in der Fremdsprache Englisch geprüft: Hörverstehen, Leseverstehen, Sprachbewusstheit, kommunikative Schreibfertigkeit, Sprechfertigkeit und interkulturelle Kompetenzen. Zusätzlich soll der generelle Sprachstand der Schülerinnen und Schüler ganzheitlich erfasst werden. Dazu wird der C-Test eingesetzt. Der C-Test ist ein gut erforschtes Testinstrument, mit dessen Hilfe auf die allgemeine Kompetenz in der Fremdsprache rückgeschlossen werden kann (vgl. Grotjahn 2002).

### **Messkonzept**

Ein C-Test besteht aus kurzen, in sich geschlossenen Texten, in denen zweite Worthälften nach bestimmten Prinzipien getilgt werden, wobei der erste und letzte Satz jeweils unverändert bleiben. Der Test basiert auf dem Prinzip der reduzierten Redundanzen: Im natürlichen Sprachgebrauch tragen verschiedene Aspekte der Sprache dazu bei, die Bedeutung einer Mitteilung zu konstituieren, so etwa Wortbedeutungen, syntaktische Aspekte oder Kollokationen. Die Redundanzen sichern die sprachliche Verständigung. Je kompetenter ein Sprachnutzer ist, desto weniger Redundanzen werden zur Sprachverarbeitung benötigt (vgl. Coleman 1996). Das integrative C-Test-Format wird dem bei der Sprachverarbeitung notwendigen komplexen Zusammenspiel von sprachlichen Fertigkeiten, kognitiven und metakognitiven Strategien und Wissensbeständen (sprachliches, strategisches und textuelles Wissen, Weltwissen) gerecht: Erfasst wird der Grad der Fähigkeit, Wortschatz-, Grammatik- und textuelles Wissen verschränkt anzuwenden: „A C-test measures the ability to apply and integrate contextual, semantic, syntactic, morphological, lexical and orthographic information and knowledge pertaining to a particular written language“ (Hastings 2002, S. 66). Den Bezug zur Realität sehen Grotjahn/Klein-Braley/Raatz (1992, S. 6) wie folgt:

„Die Notwendigkeit der Rekonstruktion gestörter Sprache kommt in der realen Welt ziemlich oft vor, z.B. am Telefon, wenn die Leitung schlecht ist, oder am Bahnhof bei Lautsprecherdurchsagen. Auch bei geschriebenen Texten findet man das Phänomen des unvollständigen bzw. beschädigten Textes, z.B. durch schlechtes Fotokopieren.“

Das Testkonstrukt eines C-Tests ist nach Sigott (2004) hoch komplex: Grundsätzlich umfasst es neben generellem sprachlichen Wissen textcharakteristisches Wissen um

Kohäsion und Kohärenz. Darüber hinaus wird die Fähigkeit erfasst, Sprache auf allen Ebenen ("from the individual letter to the text", S. 189) zu verarbeiten. Da es sich beim C-Test um Texte handelt, die rekonstruiert werden müssen, werden auch Lesestrategien und die Interpolationsfähigkeit erfasst. Letztere bezieht sich auf die Fähigkeit, einen Text vorwärts wie rückwärts zu lesen, um dessen Bedeutung zu (re)konstruieren. Hohe Reliabilität und hinreichende Validität dieses Testformats haben sich in zahlreichen Untersuchungen erwiesen (vgl. etwa Klein-Braley 1985; Raatz/Klein-Braley 1985; Grotjahn 1987 u.a.).

Das dem C-Test im DESI-Projekt zugrunde gelegte Kompetenzkonzept geht davon aus, dass umso mehr und umso anspruchsvollere Lücken gefüllt werden können, je kompetenter die Schülerinnen und Schüler in der Fremdsprache sind: Durch Wortelisionen reduzierte sprachliche Redundanzen können durch antizipatorische Sprachverarbeitung aufgelöst und die fehlenden Worthälften mittels sprachlichen und außersprachlichen Wissens rekonstruiert werden. Je automatisierter die Sprachverarbeitung abläuft, desto schneller dürften die Lücken zu füllen sein. (Vgl. Lehmann/Gänsfuß/Peek 1999). Daher kann der C-Test als Instrument zur Messung des globalen Sprachstands eingesetzt werden.

Sigott (2004, S. 189) weist auf den Zusammenhang zwischen Probandenfähigkeit und Testschwierigkeit hin: "[T]he extent to which the[se] different aspects of the construct are tapped by the individual C-Test passage is likely to be a function of personal ability and passage difficulty, and therefore not easily predictable." Schwächere Lerner können oft assoziative Ergänzungen vornehmen (Welches Wort fällt mir als erstes zu „ju...“ ein?) oder Lücken aus dem direkten Kontext ergänzen. Es ist davon auszugehen, dass die vorhandenen Sprachverarbeitungskapazitäten mit Lese- und Dekodierprozessen ausgelastet sind, so dass für anspruchsvollere Prozesse wenig bis keine Kapazitäten frei sind. Darüber hinaus sind anspruchsvollere sprachliche Phänomene (noch) nicht vorhanden. Nach Sigott (ebd.) benötigen schwächere Lerner mehr Kontext, um fehlendes sprachliches Wissen und noch nicht automatisierte Sprachverarbeitungsfähigkeiten zu kompensieren. Je fortgeschrittener die Lerner sind, desto weniger Kontext werde benötigt. Stärkere Lerner verfügen vermutlich über mehr freie Kapazitäten durch ausgeprägtere Wissensbestände und die Automatisierung bestimmter Prozesse, so dass anspruchsvollere Lücken in schwierigen Texten gelöst werden können. Die Schwierigkeit ergibt sich unter anderem dadurch, dass „... schwierigere Texte (...) seltenere und längere Wörter und längere Sätze enthalten, die an die Verarbeitungskapazität, das Weltwissen, die Fähigkeit zu schlussfolgerndem Denken usw. andere Ansprüche stellen als leichtere Texte“ (vgl. Freese 1994, S. 310). Daneben tragen die Thematik der Texte und ihre Struktur zur Schwierigkeit bei. Die Lücken selbst stellen unterschiedliche sprachliche Ansprüche, die die Schwierigkeit mitbestimmen. Auch die Anzahl der möglichen Schließungsalternativen (vgl. Stemmer 1991) ist schwierigkeitsbestimmend: Wenn nämlich der gesamttextuelle Zusammenhang genutzt werden muss, um den semantischen Kern einer Lücke zu erkennen, spielen andere psychologische Vorgänge eine Rolle als wenn dieser semantische Kern intuitiv gefasst werden kann. Die

Schwierigkeit eines C-Tests wird nun einmal bestimmt durch den Zusammenhang zwischen Text und Lücken; auch die Platzierung der Lücken bestimmt das benötigte Maß von Lese- und Sinnerschließungsstrategien.

Das Schaubild in Abbildung 1 verdeutlicht die Zusammenhänge zwischen den Kompetenzen, die bei dieser Testform eingesetzt werden müssen, und den Anforderungen, die an die Probanden gestellt werden.

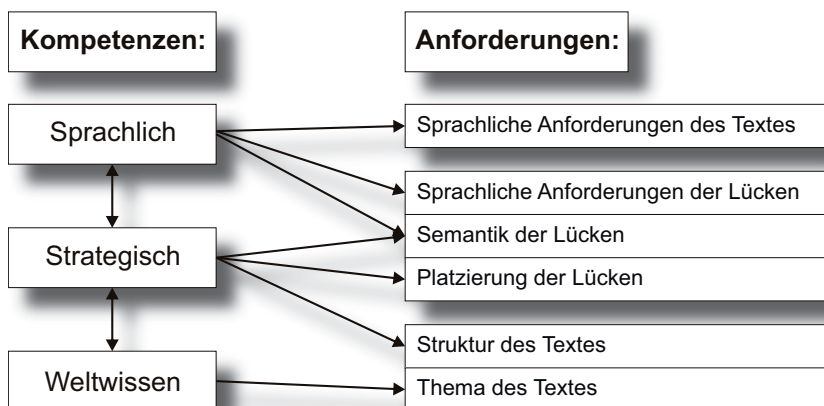


Abbildung 1: Messkonzept des C-Tests in DESI.

Die sprachliche Kompetenz umfasst orthographisches, lexikalisches, semantisches, morphologisches, syntaktisches, grammatisches, pragmatisches und textuelles Wissen (vgl. auch Bachman/Palmer 1996). Die strategische Kompetenz bezieht sich auf Prozesse des Texterschließens, auf Lesestrategien und Interpolation. Die durch einen C-Test gestellten Anforderungen werden im Abschnitt *Aufgabenmerkmale* beschrieben.

Die C-Tests des DESI-Projekts wurden im Hinblick auf die 9. Klassen des deutschen Schulsystems entwickelt: Sie basieren in ihren linguistischen und textuellen Anforderungen auf Curriculumsanalysen, sind auf den Horizont der Zielgruppe hin abgestimmt und erfassen eine möglichst große Bandbreite von Themen. Eingesetzt wurden zwölf didaktisierte Kurztex te, in denen nach einem modifizierten Tilgungsprinzip (vgl. dazu auch Arras/Eckes/Grotjahn 2002) die zweite Hälfte jedes dritten Wortes elidiert wurde; Ausnahmen davon wurden gemacht, um die Trennschärfe zu erhöhen; auch wenn es sich bei dem zu tilgenden Wort um Eigennamen oder um Wiederholungen handelte, wurde vom Tilgungsprinzip abgewichen: Wenn dasselbe Lexem (beispielsweise das hochfrequente Lexem *the*) aufeinander folgend getilgt werden sollte, wurde stattdessen das nächstfolgende Wort getilgt. Insgesamt hat jeder Text 25 Lücken. Die korrekte Lösung und mögliche Alternativen wurden vorab durch *native speakers* validiert. Zur Erzielung angemessener Trennschärfe wurden Orthographiefehler als nicht korrekt gewertet (vgl. dazu auch Arras/Eckes/Grotjahn 2002). Die C-Tests wurden in Präpilotierungen und in der Pilotierung aller DESI-Instrumente auf ihre Güte hin geprüft. Auf Basis der in der Pilotierung ermittelten Schwierigkeiten wurden einige C-Tests ausgewählt, die nur in bestimmten Schulformen eingesetzt wur-



den (z.B. leichte Tests für die Hauptschule), zusätzlich wurden zur Verankerung schulformübergreifende C-Tests verwendet. Dadurch wird eine schulformübergreifende Skalierung ermöglicht. Jede Schülerin und jeder Schüler bearbeitete zu beiden Testzeitpunkten jeweils vier Texte mit insgesamt 100 Lücken, für die 20 Minuten zur Verfügung standen.

## Aufgabenmerkmale

Um dem erwähnten Zusammenhang zwischen Probandenfähigkeit und Testschwierigkeit Rechnung zu tragen, wurden a priori Aufgabenmerkmale aus den gerade genannten Testanforderungen entwickelt. Im Rahmen des DESI-Projekts ist es jedoch aus zeitlichen und finanziellen Gründen nicht möglich, diese Merkmale empirisch zu validieren. Die Beschreibung der Aufgabenmerkmale soll helfen, die Tests in ihren Anforderungen und Schwierigkeiten zu charakterisieren. Zugleich sind diese Merkmale bei der Schwellenfindung der Kompetenzskalen dienlich (vgl. dazu den Abschnitt Kompetenzskala).

Der C-Test ist hinsichtlich seiner Merkmale auf zwei Ebenen zu beschreiben: Einmal müssen die *Texte*, die als Testgrundlage dienen, in ihren Merkmalen gesichtet werden, und zum anderen müssen die *Lücken*, d.h. die einzelnen Items, in ihren charakteristischen Merkmalen beschrieben werden.

## Aufgabenbeschreibung auf Textebene

### Thema

Die Thematik eines Textes wird hier in drei Klassen eingeteilt: *konkret*, *komplex* und *abstrakt*. Ein *konkretes* Thema oder ein Thema aus der Lebenswelt der Jugendlichen, das ihnen unmittelbar zugänglich ist, ist leichter zu verstehen und zu rekonstruieren als eine *komplexere* Darstellung oder eine Darstellung, die nicht unmittelbar im Zentrum der Lebenswelt der Schüler steht, wenn sie auch ohne Spezialwissen zu erfassen ist. Eine *abstrakte* Thematik, die den Lernenden ferner steht und zu deren Verständnis auch außersprachliches Wissen aktiviert werden muss, ist am schwierigsten zu erfassen. Das Merkmal „Thema“ wird wie folgt eingestuft:

- 0 direkter Zugang: aufgrund konkreter Darstellung und bekannter Textstrukturen Aktivierung bekannter Schemata;
- 1 inferierender Zugang: kein Spezialwissen nötig, doch müssen z.B. Strukturen oder Sachverhalte interpretiert, schlussgefolgert werden; bestimmte Abschnitte verlangen mentale Modelle; Darstellung mittlerer Komplexität oder Abstraktheit;
- 2 rekonstruktiver Zugang: teils Spezialwissen erforderlich; mentale Modellbildung zum Verständnis notwendig; Strukturen oder Sachverhalte müssen teils rekonstruiert, teils ergänzt werden; abstrakte Darstellung; vertraute Schemata im Regelfall nicht dazu vorhanden.

## Textniveau

Das Merkmal „Textniveau“ bezieht sich auf die sprachlichen Anforderungen an den Text, auf den lexikalischen und grammatischen Schwierigkeitsgrad der Texte. Die Einschätzung der Schwierigkeiten nach Klein-Braleys Vorgehen über die Nutzung der *type-token ratio* und der durchschnittlichen Satzlänge (vgl. Klein-Braley 1985) zeigte keinen signifikanten Zusammenhang mit dem Schwierigkeitsgrad der Tests. Deshalb wurde folgende, an den Curricula orientierte Schwierigkeitsabstufung hinsichtlich des Textniveaus vorgenommen:

*Leicht:* Allgemeinsprachlicher Wortschatz aus dem curricularen Kernbereich; parataktische Strukturen; einfache Satzverknüpfungen; weitgehender Verzicht auf komplexe Zeiten und Zeitgefüge sowie auf Passivformen.

*Mittel:* Wortschatz nicht nur aus dem curricularen Kernbereich, doch weitgehender Verzicht auf spezialisiertes Vokabular; para- und hypotaktische Strukturen; Satzverknüpfungen, die über die einfachsten Mittel hinausgehen; gebräuchliche komplexe Zeiten und Zeitgefüge; frequente Passivformen.

*Schwer:* Spezialisiertes Vokabular; komplexe Satzstrukturen; alle Arten von Satzverknüpfungen; keine Einschränkung im Bereich der grammatikalischen Phänomene.

Allerdings lässt diese Einschätzung allein keineswegs direkte Rückschlüsse auf die Schwierigkeit des Tests zu, denn dabei spielen weitere Faktoren wie die Thematik (siehe oben), die benötigten Lösestrategien oder die Schwierigkeit der individuellen Lücken eine Rolle. Das Merkmal „Textniveau“ wird wie folgt kodiert:

- 0 leicht: Wortschatz – Strukturwörter, curriculärer Kernbereich; Grammatik – Basisstrukturen (Parataxe, Verzicht auf komplexe Strukturen);
- 1 mittelschwer: Wortschatz – auch jenseits Grundwortschatz, doch kein Spezialvokabular; Grammatik – gebräuchliche und frequente Strukturen und Phänomene;
- 2 fortgeschritten: Wortschatz – auch Spezialvokabular, Kollokationen, Idiomatik; Grammatik – komplexe Strukturen, ungewöhnliche Phänomene.

## Lösestrategien

Das Merkmal „Lösestrategien“ bezieht sich auf die Komplexität des Lösungsprozesses und der dabei anzuwendenden Strategien. Auch wenn der C-Test überwiegend im Bereich der Reproduktion (automatisierte Prozesse) und der Verknüpfung von Text und Kontext angesiedelt ist, so spielen Wissenstransfer und Interpolation doch eine nicht zu vernachlässigende Rolle (vgl. auch Stemmer 1991).

Unter „Lösestrategien“ verstehen wir einerseits den Rekurs auf jenes *Weltwissen*, das benötigt wird, um eine Lücke zu schließen, und andererseits den Einsatz von *Interpolationstechniken*: Sie ermöglichen es, eine Lücke unter Zuhilfenahme aller im Text gegebenen Informationen zu schließen. Items, die eine Reproduktion von deklarativem Wissen oder das Abspulen von Routinen verlangen, dürften einfacher zu lösen sein also solche, die eine Verknüpfung von Fakten, textuellen Informationen und Wissensbeständen voraussetzen; am schwierigsten dürften Items zu lösen sein,

die Verallgemeinerungen und Transfer von Wissen, Fakten oder Informationen erfordern. Dieses Merkmal ermöglicht es, bestimmte Fertigkeiten im Zusammenhang mit Lesekompetenzen und Texterschließungskompetenzen zu erfassen, es lässt aber auch Rückschlüsse auf die Textschwierigkeit zu. Das Merkmal „Lösestrategien“ wird wie folgt eingestuft:

- 0 kein (oder nur beschränktester) Rekurs auf Weltwissen erforderlich, keine Interpolationstechniken nötig;
- 1 geringer Rekurs auf Weltwissen und geringe Nutzung von Interpolationstechniken notwendig;
- 2 mittlerer Bedarf an Weltwissen und Interpolationstechniken;
- 3 starker Bedarf an Weltwissen und hoher Einsatz von Interpolationstechniken notwendig.

## **Aufgabenbeschreibung auf Itemebene** (bezogen auf die Lücken eines Textes)

### **Sprachniveau**

Der Begriff „Sprachniveau“ bezeichnet die sprachlichen Anforderungen an die Items, das Schwierigkeitsniveau der sprachlichen Phänomene, die mit den jeweiligen Lücken erfasst werden sollen. Nicht unter das Merkmal „Sprachniveau“ fällt die Schwierigkeit, die z.B. dadurch entsteht, dass zur Schließung komplexe linguistische wie rekursive Prozesse aktiviert werden müssen. Diese Art der Schwierigkeit kann nicht auf Itemebene, wohl aber auf Textebene über das Merkmal „Lösestrategien“ (siehe oben) erfasst werden. Die jeweilige Lücke wird hinsichtlich des Niveaus ihrer sprachlichen Phänomene auf drei Schwierigkeitsstufen eingeschätzt:

- 0 leicht: Wortschatz des curricularen Kernbereichs (Grundwortschatz); grammatikalische und textuelle Basisphänomene;
- 1 mittelschwer: Aufbauwortschatz; grammatikalische und textuelle Phänomene mittleren Schwierigkeitsniveaus;
- 2 fortgeschritten: Spezialwortschatz und komplexe grammatikalische wie textuelle Phänomene gehobenen Schwierigkeitsniveaus.

### **Semantik**

Bei dem Merkmal „Semantik“ soll eingeschätzt werden, ob die Bedeutung des gesuchten Lexems, seine Semantik, direkt zugänglich ist oder ob sie erschlossen oder unter Zuhilfenahme eines mentalen Modells konstruiert werden muss. Im einfachsten Fall ergibt sich die Bedeutung assoziativ aus den angegebenen Buchstaben, so dass weder Kontext noch weiterführende Überlegungen zu Hilfe genommen werden müssen (etwa „I loved my rabbit; her na... was Gillian“). Schwieriger dürften Lücken sein, deren Bedeutung aus dem Kontext erschlossen werden muss (beispielsweise „It ne... sounds right when I use an American word“). Am schwierigsten sind Lücken zu schließen, deren Semantik erst konstruiert werden muss, sei es unter Zuhilfenahme aller textuellen Informationen, die zur korrekten Inferenz führen können, oder sei es

durch den Einsatz von Weltwissen oder mentalen Modellen (etwa “A fi... broke out in the hotel and nearly killed all tourists“: Hier fand sich in der Pilotierung häufig die inkorrekte Lösung *fish* statt der gesuchten Lösung *fire* – um auf die korrekte Lösung zu kommen, muss vermutlich erst der größere semantische Kontext konstruiert werden und ein Modell der Vorgänge im Hotel gebildet werden). Die Semantik der Items wird auf drei Schwierigkeitsstufen eingeschätzt:

- 0 Semantik der Lücke direkt zugänglich, meist assoziativ;
- 1 Semantik muss erschlossen werden aus Kotext oder Kontext;
- 2 Schwierige semantische Zusammenhänge: Bedeutung des Lexems muss konstruiert werden.

### Itemfokus

Der „Itemfokus“ dient der Abschätzung, auf welche sprachlichen Teildimensionen die betreffende Lücke „vorwiegend“ fokussiert. Damit wird versucht, den C-Test als integrativen Test in seinen Einzeldimensionen besser zu fassen. Auf diese Weise entsteht „angereichertes Material“: Gruppierungen von Items, die annähernd die gleichen Teilfertigkeiten messen. Insofern ist eine Aussage, welche Teildimensionen die Lücke tatsächlich misst, nicht eindeutig möglich, denn dazu müsste man zusätzlich die Prozesse im Kopf der Schüler erfassen und die aufgetretenen Fehler analysieren. Dies ist jedoch im Rahmen des DESI-Projekts wie gesagt nicht möglich.

Es wird einerseits eingeschätzt, ob die Items vorwiegend auf lexikalische, grammatische oder textstrukturierende Phänomene fokussieren, oder ob sie auf komplexe Konfigurationen fokussieren: Letztere erfordern komplexe Operationen seitens der Probanden, wenn beispielsweise die Bedeutung des gesuchten Lexems erschlossen werden und gleichzeitig eine grammatische Endung – etwa eine Zeitform, die sich aus dem Kontext ergibt – angefügt werden muss.

Andererseits wird erfasst, ob zur Schließung einer Lücke der unmittelbare Kotext oder der textuelle wie außertextliche Kontext genutzt werden muss, oder ob es zur Schließung komplexer Informationsverarbeitung bedarf. Wenn eine Lücke nur mit Hilfe des unmittelbaren Kotexts zu schließen ist, so dürfte dies leichter fallen als wenn es zur Schließung der Zuhilfenahme eines größeren Kontexts bedarf (Assoziationen vs. kurzfristige Informationsspeicherung und -verarbeitung). Am schwersten dürfte die Schließung einer Lücke fallen, wenn dazu komplexe Informationsverarbeitung wie beispielsweise die Nutzung anderer, von der gesuchten Lücke abhängiger Items nötig ist. Es wird vermutet, dass textstrukturierende Phänomene nicht kotextuell zu schließen sind, da sie sich immer aus dem (Kon-)Text ergeben. Das Merkmal „Itemfokus“ wird wie folgt eingestuft:

- 0 Kotextuell zu schließende Lücken, die auf lexikalische oder grammatische Phänomene fokussieren;
- 1 Kontextuell oder textuell zu schließende Lücken, die auf Lexik, Grammatik oder auf Phänomene der Textstrukturierung fokussieren;

- 2 Lücken, zu deren Schließung es komplexer Informationsverarbeitung bedarf (etwa weil sie von anderen Items abhängig sind); Items, die auf komplexe linguistische wie außersprachliche Konfigurationen fokussieren.

Die schwierigkeitsbestimmenden Aufgabenmerkmale wurden a priori durch das Augsburger DESI-Team eingeschätzt; Divergenzen wurden diskutiert. Analog zu den anderen DESI-Modulen wird darüber hinaus zurzeit beim Modul C-Test untersucht, ob das hier beschriebene System der Aufgabenmerkmale von Personen, die weder die Aufgabenmerkmale noch die einzuschätzenden Aufgaben entwickelt haben, sinnvoll und vergleichbar angewandt werden kann. An dieser Validierungsstudie nehmen Mitarbeiter des DPC Hamburg teil.

## Kompetenzskala: Niveaus, Schwellen und Beispielitems

Die Items der eingesetzten C-Tests wurden Rasch-skaliert. Die so gewonnene Skala wurde in Kompetenzniveaus eingeteilt, um auf probabilistischer Basis eine inhaltliche Rückmeldung über die im C-Test erreichten Leistungen geben zu können. Diese Niveaus beschreiben inhaltlich, über welche Kompetenzen die Schüler und Schülerinnen auf dem jeweiligen Niveau verfügen.

Die Niveaueinteilungen gründen einerseits in den im vorangegangenen Abschnitt beschriebenen schwierigkeitsbestimmenden Merkmalen auf Text- und Itemebene. Mittels Regressionsanalysen wurden die Merkmale identifiziert, die am besten zur Vorhersage der Schwierigkeiten geeignet sind (vgl. Hartig in diesem Band). Die in das Regressionsmodell eingeflossenen Merkmale sind die folgenden: auf Itemebene die Merkmale „Sprachniveau“ (im Folgenden SN abgekürzt), „Semantik“ (SE) und „Itemfokus“ (IF); auf Textebene das Merkmal „Textniveau“ (TN). In Tabelle 1 sind die Regressionsgewichte dieser genannten Aufgabenmerkmale bei der Vorhersage der Aufgabenschwierigkeiten aufgelistet. Das Regressionsmodell erklärt mit  $R^2 = .51$  etwas mehr als die Hälfte der beobachteten Unterschiede zwischen den Aufgabenschwierigkeiten.

Die aus dem Regressionsmodell abgeleiteten erwarteten Schwierigkeiten bestimmter Merkmalskombinationen wurden benutzt, um Schwellen zwischen den Kompetenzniveaus zu definieren. Hierbei wurden Merkmalskombinationen gewählt, die bei vielen der C-Test Lücken tatsächlich realisiert sind.

Andererseits gründen die Niveaueinteilungen in inhaltlichen Itemanalysen: Die Items wurden nach der Skalierung in ihren Anforderungen und Schwierigkeiten beschrieben, um sie in ihren Charakteristika auch von empirischer Seite zu erfassen. Diese Beschreibungen wurden auch für die endgültige Schwellenziehung herangezogen: Die Schwellen wurden bei den erwarteten Schwierigkeiten vielfach realisierter Merkmalskombinationen gesetzt, bei denen die inhaltlichen Analysen eine solche Schwelle ebenfalls nahe legten.

*Tabelle 1: Regressionsgewichte bei der Vorhersage der Aufgabenschwierigkeiten mit den Aufgabenmerkmalen Sprachniveau (SN), Semantik (SE), Itemfokus (IF) und Textniveau (TN) ( $R^2 = .51$ ).*

<b>Merkmale</b>	<b>Regressions- gewicht</b>	<b>Standard- fehler</b>	<b>Irrtums- wahrschein- lichkeit</b>
Regressionskonstante	-0.77	0.18	< .001
<b>SN=1</b> Sprachniveau der Lücke mittelschwer	1.02	0.15	< .001
<b>SN=2</b> Sprachniveau der Lücke fortgeschritten	2.41	0.30	< .001
<b>SE=1</b> Semantik der Lücke muss erschlossen werden	0.76	0.16	< .001
<b>SE=2</b> Schwierige semantische Zusammenhänge der Lücke	1.74	0.32	< .001
<b>IF=1</b> Itemfokus: Kontextuell oder textuell zu schließende Lücken	0.54	0.16	.001
<b>IF=2</b> Itemfokus: Komplexe Informationsverarbeitung	1.12	0.24	< .001
<b>TN=1</b> Textniveau mittelschwer	0.45	0.19	.022
<b>TN=2</b> Textniveau fortgeschritten	0.67	0.21	.001

Die Kompetenzskala zum C-Test Englisch besteht aus fünf Niveaus (kodiert *KN A* bis *KN E*). Das unterste Niveau wurde nicht aufgrund des Regressionsmodells, sondern aufgrund zusätzlicher inhaltlich-curricularer Erwägungen bestimmt. Kompetenzniveau A kann auch als „Lernstufe“ betitelt werden: Die durch die dort angesiedelten Items elizitierten Leistungen sind beschreibbar, doch weit unter dem für Lernende der Klassenstufe 9 erwartbaren Niveau. Die tatsächlichen Schwierigkeiten dieser Items liegen noch unterhalb der aus dem Regressionsmodell resultierenden erwarteten Schwierigkeit für die einfachsten Aufgaben.

Auf der Kompetenzskala können einerseits die Items in ihren Schwierigkeiten und Anforderungen eingeordnet werden, und andererseits die Leistungen der Schülerinnen und Schüler, so dass sie gemäß ihrer Fähigkeiten auf einem bestimmten Niveau verortet werden können.

Tabelle 2 gibt eine Übersicht über die Kompetenzniveaus des Moduls C-Test im DESI-Projekt, die Schwellenwerte (die erwarteten Schwierigkeitswerte der Items aus dem Regressionsmodell) und die jeweils entscheidenden Merkmalskombinationen der vier ins Regressionsmodell eingeflossenen Merkmale:

*Tabelle 2: Schwellenwerte und Merkmalskombinationen der Kompetenzniveaus im C-Test. Die schattierten Felder zeigen an, auf welchem Niveau eine Merkmalsausprägung erstmalig auftritt.*

Niveau	Logit	SN	SE	IF	TN
<b>KN A</b>	-1.75	0	0	0	0
<b>KN B</b>	-0.77	0	0	0	0
<b>KN C</b>	-0.10	0	0	0	2
<b>KN D</b>	1.20	0	1	1	2
<b>KN E</b>	2.22	1	1	1	2

#### Legende

Logit: Erwarteter Schwierigkeitswert der Aufgaben mit der angegebenen Merkmalsausprägung

**SN** Items: Sprachniveau

**SE** Items: Semantik

**IF** Items: Itemfokus: ko-, kontextuell, komplex

**TN** Textniveau

Im Folgenden werden die Kompetenzniveaus und ihre Deskriptoren vorgestellt. Diese beschreiben zum einen den konkreten Umgang der Probanden mit der Aufgabenstellung; zum anderen werden generalisierende Aussagen bezüglich der zugrunde liegenden Fähigkeiten getroffen. Die Deskriptoren wurden wie folgt entwickelt:

Die in die oben erwähnte Regressionsanalyse eingeflossenen Merkmalskombinationen, die nicht in die Regressionsanalyse eingeflossenen schwierigkeitsbestimmenden Merkmale (a priori eingeschätzt) und die inhaltliche Itemanalyse (nach dem ersten Messzeitpunkt) ergeben zusammen die Beschreibung des Umgangs der Probanden mit der konkreten Aufgabenstellung. Zusätzlich werden die zugrunde liegenden Kompetenzen im Bereich des sprachlichen Wissens, der Lesestrategien und der Fähigkeiten der Textrezeption beschrieben. Die in diesem Zusammenhang gewählten Formulierungen lehnen sich, soweit es sinnvoll erscheint, an die can-do-Formulierungen des Gemeinsamen Europäischen Referenzrahmens (GER) an.<sup>1</sup>

<sup>1</sup> Das Testkonstrukt des C-Tests kann als Ganzes nicht im GER verankert werden. Dies ist so angesichts der ganzheitlichen Natur des C-Tests und der Konzeption des Referenzrahmens. Dennoch lassen sich Berührungspunkte bei folgenden Kategorien des GER ausmachen: Lesen (74f), Rezeptionsstrategien (78), Texte verarbeiten (98), Lexik (112f), Grammatik (114), Orthographie (118), Kohäsionsmittel (125), Spektrum allgemein (110). Dabei wurde der „Wortlaut der Deskriptoren“ der GER-Skalen mit den Beschreibungen der DESI-Niveaus verglichen (vgl. dazu auch Europarat 2001, S. 34 und 203). Die DESI-Niveaus A mit E zeigen Berührungspunkte mit den GER-Niveaus A1 mit C1, doch im DESI-Projekt kann eine empirische Anbindung, wie sie etwa das *Manual* zum GER (Council of Europe 2003) vorschlägt, nicht erfolgen. Deshalb handelt es sich bei dem hier erwähnten Abgleich lediglich um eine erste qualitative Ein-



Jedes Niveau wird durch ein Beispielitem illustriert, das die für dieses Niveau typischen Merkmale trägt und dessen empirisch ermittelte Schwierigkeit der durch das Regressionsmodell erwarteten Schwierigkeit für diese Merkmalskombination entspricht.

### Kompetenzniveau A

- Kann im C-Test vorwiegend Lücken des sprachlichen Basisbereichs assoziativ oder unter Zuhilfenahme des unmittelbaren Kontexts schließen, wobei es sich überwiegend um vertraute Phänomene handelt, die hochfrequent sind oder Parallelen zur Muttersprache aufweisen. Beispielsweise können einfache Strukturwörter oder einfachste, sich aus dem Kontext aufdrängende Inhaltswörter ergänzt werden.
- Verfügt über einen elementaren Vorrat von Wörtern und Basisstrukturen in einem teils auswendig gelernten Repertoire, die sich auf konkrete Situationen beziehen. Kann basales textuelles Wissen anwenden (einfachste Konnektoren wie *and*, *but*).

**Beispielitem** [ $\sigma = -1.21$  resp.  $-1.33$  (Schwelle:  $-1.75$ )]<sup>2</sup>:

*[The rabbit] sat un\_\_\_\_\_ a tree in our gar\_\_\_\_\_.*

Hier muss eine dem Deutschen ähnliche Präposition respektive ein dem Deutschen ähnliches Nomen ergänzt werden in einer bekannten Struktur; beide lexikalischen Items sind Teil des Grundwortschatzes; die Semantik ist unmittelbar zugänglich. Die Items können mittels eindeutigen Kontexts (*tree*) geschlossen werden und befinden sich in einem leichten Text.

### Kompetenzniveau B

- Kann im C-Test bei leichten Texten solche Lücken schließen, die sich auf sprachliche Phänomene des curricularen Kernbereichs beziehen und deren Semantik unmittelbar zugänglich ist. Kann in mittelschweren Texten die leichten Items lösen. Zu den leichten Items zählen einfache Inhalts- und frequente Strukturwörter oder die Ausführung einfacherer, frequenter grammatischer Operationen an den Lücken.
- Beherrscht einen begrenzten, hochfrequenten Wortschatz der Alltagssprache sowie ein begrenztes, gebräuchliches Spektrum an grammatischen Mitteln, die sich auf Routinesituationen beziehen. Kann die häufigsten Textkonnektoren wie *there*, *then*, *because* verwenden.
- Kann kurze, einfache Texte rezipieren und zum Teil rekonstruieren, wenn sie sich in gängiger, alltagsbezogener Sprache auf alltägliche Themen im Rahmen der eigenen Erfahrungen beziehen und relativ konkret in der Darstellung bleiben.

---

schätzung. Die DESI-Konsorten erheben daher auch nicht den Anspruch, die DESI-C-Test-Niveaus an die Niveaus des GER anzubinden.

2 In diesem Fall wurden zwei Beispielitems ausgewählt, da sie das gleiche Phänomen illustrieren und im betreffenden Text hintereinander vorkommen.



**Beispielitem** [Merkmale 0000<sup>3</sup>,  $\sigma = -0.65$  (erwartet:  $-0.77$  = Schwelle)]:

*[She] waited patiently for me t \_\_\_\_\_ come (...).*

Hier muss in einer nicht hochfrequenten Struktur das Infinitiv-Partikel *to* ergänzt werden; das Item ist dem sprachlichen Basisbereich zuzuordnen; die Semantik ist eindeutig; das Item kann aus dem Kotext gelöst werden und befindet sich in einem leichten Text.

### Kompetenzniveau C

- Kann im C-Test bei leichten Texten sprachlich mittelschwere Items ergänzen, wenn deren Semantik unter Zuhilfenahme des Ko- und Kontexts erschließbar ist. Kann in mittelschweren Texten sprachlich leichte Items lösen, wenn deren Semantik aus dem Ko-/Kontext zu erschließen ist. Kann in mittelschweren Texten sprachlich mittelschwere Items lösen, die kotextuell zu erschließen sind. Kann in schweren Texten sprachlich einfache Items aus dem unmittelbaren Kotext lösen, wenn deren Semantik entweder direkt/assoziativ zugänglich ist oder aus dem Ko-/Kontext erschließbar ist. Kann, soweit erforderlich, gängige grammatische Operationen an den Lücken ausführen.
- Zeigt eine gute Beherrschung des Grundwortschatzes; verfügt über ein Repertoire häufiger Floskeln und Wendungen; zeigt gute Beherrschung der gängigen grammatischen Strukturen. Verfügt über gängige, meist lineare Kohäsionsmittel. Allerdings kommt es bei Unvertrautem zu Fehlern.
- Kann allem Anschein nach Interpolationstechniken und Weltwissen nutzen, um einige schwierigere Lücken zu erschließen. Kann den Sinn des rekonstruierten Textes ansatzweise überprüfen.
- Kann unkomplizierte fiktive Texte oder Sachtexte aus dem eigenen Interessengebiet oder einem erschließbaren Gebiet rezipieren und rekonstruieren, wobei sich die Darstellungen auf einem mittleren Abstraktheitsgrad bewegen.

**Beispielitem** [Merkmale 0002,  $\sigma = -0.03$  (erwartet:  $-0.10$  = Schwelle)]:

*(...) even t \_\_\_\_\_ strictest chauvinist or macho has to accept the fact [that there are women in almost every trade and profession (...).]*

Hier muss ein gängiges grammatisches Element, der Artikel *the*, nach einem emphatischen Lexem ergänzt werden, wobei es nur diese eine korrekte Alternative gibt. Das zu ergänzende Item ist dem sprachlichen Basisbereich zuzuordnen. Die Semantik ist unmittelbar zugänglich. Das Item befindet sich in einem schweren Text.

### Kompetenzniveau D

- Kann im C-Test bei mittelschweren Texten sprachlich mittelschwere Items ergänzen, deren Semantik unter Zuhilfenahme des Ko-/Kontexts erschlossen werden muss. Kann in schweren Texten Items des sprachlichen Basisbereichs und des mittleren sprachlichen Niveaus lösen, deren Semantik unter Zuhilfenahme des Ko-/Kontexts zu erschließen ist. Kann, soweit erforderlich, an den Lücken gram-

3 Diese Angaben zu den Merkmalsausprägungen der Beispielitems beziehen sich in immer gleicher Reihenfolge auf die o.g. Merkmale *Sprachniveau der Items*, *Semantik der Items*, *Itemfokus*, *Textniveau*.

matische Operationen, bis auf wenige seltene, ausführen. Kann einige komplexe, voneinander abhängige Lücken schließen.

- Verfügt über einen großen Wortschatz zu den meisten allgemeinen Themen und zu eigenen Interessengebieten. Verfügt über eine gute Beherrschung auch komplexerer grammatischer Phänomene. Kann eine Anzahl verschiedener inhaltlicher wie textueller Verknüpfungsmittel hinreichend korrekt verwenden.
- Kann allem Anschein nach Interpolationstechniken und Weltwissen einsetzen, um teils schwierige semantische Beziehungen zu rekonstruieren und den Text zu einem sinnvollen Ganzen zu ergänzen. Kann Überprüfungsstrategien anwenden. Kann allem Anschein nach verschiedenste Wissensbestände integrativ nutzen.
- Kann auch komplexere fiktive und Sachtexte rezipieren und rekonstruieren, wobei deren Thema außerhalb der Lebenswelt der Lernenden liegen kann; die thematische Darstellung ist im Regelfall komplex, sie kann abstrakt sein und ansatzweise Spezialwissen verlangen.

**Beispielitem** [Merkmale 0112,  $\sigma = 1.20$  (erwartet: 1.20 = Schwelle)]:

*B \_\_\_\_\_ they had to do without [them] for two days (...).*

Das hier gesuchte Item *but* ist dem sprachlichen Basisbereich zuzuordnen. Seine kontrastive Semantik muss über den unmittelbaren Satz hinaus erschlossen werden, so dass Kotext wie Kontext genutzt werden müssen, um den Konnektor (textuelles Strukturierungsmittel) korrekt zu ergänzen. Das Item befindet sich in einem schweren Text.

### Kompetenzniveau E

- Kann im C-Test in allen Arten von Texten solche Items im mittleren sprachlichen Schwierigkeitsbereich lösen, die auf komplexe Konfigurationen fokussieren, wobei deren Semantik über den Ko-/Kontext und das Weltwissen zu erschließen ist. Kann sprachlich schwere Items ergänzen, deren Semantik zu (re-)konstruieren ist. Kann dabei alle notwendigen Informationen aus Kotext, Kontext und Weltwissen zur Schließung von komplex interagierenden Lücken nutzen; kann auch weniger frequente grammatische Operationen an den Lücken ausführen.
- Beherrscht einen umfangreichen Wortschatz, zum Teil auch zu Spezialgebieten; kann idiomatische Ausdrücke und umgangssprachliche Wendungen hinreichend korrekt verwenden. Beherrscht ein umfangreiches Spektrum an grammatischen Strukturen. Kann auch weniger frequente Mittel der Textverknüpfung einsetzen.
- Kann allem Anschein nach Weltwissen, Lesestrategien und Interpolationstechniken einsetzen zur Bedeutungerschließung auch bei anspruchsvollen semantischen Zusammenhängen. Kann den fertigen Text auf Sinnzusammenhang überprüfen.
- Kann auch komplexe, anspruchsvolle Texte verstehen und rekonstruieren; die Thematik kann der Lebenswelt eher fern stehen und teils Spezialwissen verlangen; die Darstellung kann abstrakt sein.

**Beispielitem** [Merkmale 1112,  $\sigma = 2.25$  (erwartet: 2.22 = Schwelle)]:

*The idea (...) began ea \_\_\_\_\_ in this century (...).*

Das Item ist in dieser Struktur einem mittleren Sprachniveau zuzuordnen; die Semantik lässt sich über den Kontext erschließen, wobei ein mentales Modell der Textaussage dabei helfen kann; die Lücke kann kontextuell geschlossen werden. Das Item befindet sich in einem schweren Text.

## Literatur

- Arras, U./Eckes, T./Grotjahn, R. (2002): C-Tests im Rahmen des 'Test Deutsch als Fremdsprache' (TestDaF): Erste Forschungsergebnisse. In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 4). Bochum: AKS, S. 175-209.
- Bachman, L.F./Palmer, A.S. (1996): Language testing in practice. Oxford: OUP.
- Coleman, J.A. (1996): Studying Languages. A Survey of British and European Students. London: CILT. Insbesondere "Part 2 – Research Procedures. The C-Test", S. 136-161.
- Council of Europe (2003): Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual. Preliminary Pilot Version. Strasbourg: Language Policy Division (DGIV/EDU/LANG (2003) 5 rev. 1).
- Europarat (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Berlin: Langenscheidt.
- Freese, H.-L. (1994): Was misst und was leistet 'Leistungsmessung mittels C-Tests'? In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 2). Bochum: Brockmeyer, S. 305-311.
- Grotjahn, R. (1987): How to Construct and Evaluate a C-Test: A Discussion of Some Problems and Some Statistical Analyses. In: Grotjahn, R./Klein-Braley, C./Stevenson, D.K. (eds.): Taking their Measure: The Validity and Validation of Language Tests. Bochum: Brockmeyer 1987, S. 219-253.
- Grotjahn, R. (2002): Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 4). Bochum: AKS, S. 211-225.
- Grotjahn, R./Klein-Braley, C./Raatz, U. (1992): C-Tests in der praktischen Anwendung. Erfahrungen beim Bundeswettbewerb Fremdsprachen. In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 1). Bochum: Brockmeyer, S. 1-34.
- Hastings, A.J. (2002): Error analysis of an English C-test: Evidence for integrated processing. In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendungen (Bd. 4). Bochum: AKS, S. 53-66.
- Klein-Braley, C. (1985): Advance Prediction of Test Difficulty. In: Klein-Braley, C./Raatz, U. (Hrsg.): Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis. Bochum: AKS, S. 23-41.
- Lehmann, R.H./Gänsfuß, R./Peek, R. (1999): Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 7. Hamburg: Behörde für Schule, Jugend und Berufsbildung.
- Raatz, U./Klein-Braley, C. (1985): How to Develop a C-Test. In: Klein-Braley, C./Raatz, U. (Hrsg.): Fremdsprachen und Hochschule 13/14: Thematischer Teil: C-Tests in der Praxis. Bochum: AKS, S. 20-22.
- Sigott, G. (2004): Towards Identifying the C-Test Construct. Frankfurt: Lang.
- Stemmer, B. (1991): What's on a C-test taker's mind? Mental processes in C-test taking. Bochum: Brockmeyer.